

RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/804,481A
Source: 1600
Date Processed by STIC: 11/3/2003

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03



1600

RAW SEQUENCE LISTING

DATE: 11/04/2003

PATENT APPLICATION: US/09/804,481A

TIME: 13:57:42

Input Set : N:\Crf4\11032003\I804481A.raw

Output Set : N:\CRF4\11042003\I804481A.raw

1 <110> APPLICANT: de Graaf, David
 2 Lander, Eric S.
 3 <120> TITLE OF INVENTION: Novel Small Nuclear RNA Vectors and Uses
 4 Therefor
 5 <130> FILE REFERENCE: 2825.1023-001
 C--> 6 <140> CURRENT APPLICATION NUMBER: US/09/804,481A
 C--> 7 <141> CURRENT FILING DATE: 2001-03-12
 8 <150> PRIOR APPLICATION NUMBER: 60/188,304
 9 <151> PRIOR FILING DATE: 2000-03-10
 10 <160> NUMBER OF SEQ ID NOS: 11
 11 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 13 <210> SEQ ID NO: 1
 14 <211> LENGTH: 4639
 15 <212> TYPE: DNA
 16 <213> ORGANISM: Homo sapiens
 17 <400> SEQUENCE: 1
 18 gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg 60
 19 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
 20 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgagct cggtagcccg 180
 21 ggagatccgg taaggaccag cttcttttgg agagaacaga cgcaggggcg ggagggaaaa 240
 22 agggagaggg agacgtcact tccccttggc ggctctggca gcagattggt cggttgagtg 300
 23 gcagaaaggc agacggggac tgggcaaggc actgtcggtg acatcacgga caggcgact 360
 24 tctatgtaga tgaggcagcg cagaggctgc tgcttcgcca cttgctgctt caccacgaag 420
 25 gaggttcccg gccctgggag cgggttcagg accgctgacg ggaagtgaga atcccagctg 480
 26 tgtgtcaggg ctggaaaggg ctcgggagtg cgcggggcaa gtgaccgtgt gtgtaaagag 540
 27 tgaggcgatg gaggtgtgt cggggcagag gcccaagatc tcaagggcc ataactgtg 600
 28 taccatcgat tgcaggggag ataccatgat cacgaagggt gttttccag ggcgaggtt 660
 29 atccattgca ctccgatgt gctgaccct gcgatttccc caaagcttg aaactcgact 720
 30 gcataatttg tggtagtggg ggactgcgtt cgcgctttcc ectgacttcc tggagtttca 780
 31 aaagtagact gtacgctaac cggatcctct agagtcgacc tgcaggcatg cagaagacaa 840
 32 ttagcaggca tgcctgggat gcggtgggct ctatggcttc tgaggcggaa agaaccagct 900
 33 ggggctctag ggggtatccc cagcgccct gtacggcgcc attagcgcg gcgggtgtg 960
 34 tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt 1020
 35 tcttcccttc ctttctcgcc acgttcgccc gctttccccg tcaagctcta aatcggggca 1080
 36 tcccttttag gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 1140
 37 gtgatgggtc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg 1200
 38 agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 1260
 39 cgggtctattc ttttgattta taagggattt tggggatttc ggcctattgg ttaaaaaatg 1320
 40 agctgattta acaaaaattt aacgcgaatt aattctgtgg aatgtgtgtc agtttaggtg 1380
 41 tggaaagtcc ccaggctccc caggcaggca gaagtatgca aagcatgcat ctcaattagt 1440
 42 cagcaaccag gtgtggaaag tcccaggct cccagcagg cagaagtatg caaagcatgc 1500
 43 atctcaatta gtcagcaacc atagtccgc cctaactcc gccatcccg cccctaactc 1560
 44 cgcccagttc cgcccattct ccgcccctat gctgactaat tttttttatt tatgcagagg 1620

Does Not Comply
Corrected Diskette Needed

P.3

RAW SEQUENCE LISTING

DATE: 11/04/2003

PATENT APPLICATION: US/09/804,481A

TIME: 13:57:42

Input Set : N:\Crf4\11032003\I804481A.raw

Output Set: N:\CRF4\11042003\I804481A.raw

45	ccgaggccgc	ctctgcctct	gagctattcc	agaagtagtg	aggaggcttt	tttggaggcc	1680
46	taggcttttg	caaaaagctc	ccgggagctt	gtatatccat	tttcggatct	gatcagcacg	1740
47	tgttgacaat	taatcatcgg	catagtatat	cggcatagta	taatacgaca	aggtgaggaa	1800
48	ctaaaccatg	gccaaagtga	ccagtgccgt	tccggtgctc	accgcgcgcg	acgtcgcccg	1860
49	agcggtcgag	ttctggaccg	accggctcgg	gttctcccg	gacttcgttg	aggacgactt	1920
50	cgcgggtgtg	gtccgggacg	acgtgaccct	gttcatcagc	gcggtccagg	accaggtggt	1980
51	gccggacaac	accctggcct	gggtgtgggt	gcgcggcctg	gacgagctgt	acgccgagtg	2040
52	gtcggaggtc	gtgtccacga	acttccggga	cgctccggg	ccggccatga	ccgagatcgg	2100
53	cgagcagccg	tgggggcg	agttcgccct	gcgcgaccg	gccggcaact	gcgtgcactt	2160
54	cgtggccgag	gagcaggact	gacacgtgct	acgagatttc	gattccaccg	ccgccttcta	2220
55	tgaaggttg	ggcttcggaa	tcgttttccg	ggacgccggc	tggatgatcc	tccagcgcg	2280
56	ggatctcatg	ctggagttct	tcgccaccc	caacttggtt	attgcagctt	ataatggtta	2340
57	caaataaagc	aatagcatca	caaatttcac	aaataaagca	tttttttcac	tgcatcttag	2400
58	ttgtggtttg	tccaaactca	tcaatgtatc	ttatcatgtc	tgtataccgt	cgacctctag	2460
59	ctagagcttg	gcgtaatcat	ggtcatagct	gtttcctgtg	tgaattgtt	atccgctcac	2520
60	aattccacac	aacatacgag	ccggaagcat	aaagtgtaaa	gcctggggtg	cctaattgagt	2580
61	gagctaactc	acattaattg	cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgtc	2640
62	gtgccagctg	cattaatgaa	tccggcaacg	cgcggggaga	ggcgggtttc	gtattggcg	2700
63	ctcttcgctt	tctcgcgtca	ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	2760
64	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	2820
65	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	2880
66	gtttttccat	aggctccgcc	cccctgacga	gcatacaaaa	aatcgacgct	caagtcagag	2940
67	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	gtcctcctcg	3000
68	gcgctctcct	gttccgacct	tgcgcgttac	cggatacctg	tccgccttcc	tcccttcggg	3060
69	aagcgtggcg	ctttctcaat	gctcacgctg	taggtatctc	agttcggtgt	aggtcgttcg	3120
70	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	3180
71	taactatcgt	cttgagtcga	acccggtaag	acacgactta	tcgccactgg	cagcagccac	3240
72	tggtaacagg	attagcagag	cgaggtatgt	aggcgggtgt	acagagttct	tgaagtgtg	3300
73	gcctaactac	ggctacacta	gaaggacagt	atttggtatc	tgcgctctgc	tgaagccagt	3360
74	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	3420
75	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	3480
76	tttgatcttt	tctacggggt	ctgacgctca	gtggaacgaa	aactcacgtt	aagggatttt	3540
77	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	3600
78	taaatcaatc	taaagtatat	atgagtaaac	ttggtctgac	agttaccaat	gcttaatcag	3660
79	tgaggcacct	atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	3720
80	cgtgtagata	actacgatac	gggagggtct	accatctggc	cccagtgtcg	caatgatacc	3780
81	cgagagccca	cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	ccggaaggcg	3840
82	cgaagcgaga	agtggtcctg	caactttatc	cgctccatc	cagtctatta	attgttgccg	3900
83	ggaagctaga	gtaagtagtt	cgccagttaa	tagtttgctc	aacgttggtg	ccattgctac	3960
84	aggcatcgtg	gtgtcacgct	cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	4020
85	atcaaggcga	gttacatgat	cccccatggt	gtgcaaaaaa	gcggttagct	ccttcgggtc	4080
86	tccgatcgtt	gtcagaagta	agttggccgc	agtgttatca	ctcatggtta	tggcagcact	4140
87	gcataattct	cttactgtca	tgcctccgtg	aagatgcttt	tctgtgactg	gtgagtactc	4200
88	aaccaagtca	ttctgagaat	agtgatatcg	gcgaccgagt	tgctcttgcc	cggcgtaaat	4260
89	acgggataat	accgcgccac	atagcagaac	tttaaaagt	ctcatcattg	gaaaacgttc	4320
90	ttcggggcga	aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	tgtaacccac	4380
91	tcggtcaccc	aactgatctt	cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	4440
92	aacaggaag	caaaatgccg	caaaaaagg	aataaggcg	acacggaaat	ggtgaatact	4500
93	catactcttc	ctttttcaat	attattgaag	catttatcag	ggttattgtc	tcatgagcgg	4560

RAW SEQUENCE LISTING

DATE: 11/04/2003

PATENT APPLICATION: US/09/804,481A

TIME: 13:57:42

Input Set : N:\Crf4\11032003\I804481A.raw

Output Set: N:\CRF4\11042003\I804481A.raw

```

94      atacatatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca catttccccg 4620
95      aaaagtgccca cctgacgtc                                         4639
97 <210> SEQ ID NO: 2
98 <211> LENGTH: 5
99 <212> TYPE: DNA
100 <213> ORGANISM: Artificial Sequence
101 <220> FEATURE:
102 <223> OTHER INFORMATION: single-stranded restriction fragment overhand
103 <400> SEQUENCE: 2
104      gcagg                                                         5
106 <210> SEQ ID NO: 3
107 <211> LENGTH: 5
108 <212> TYPE: DNA
109 <213> ORGANISM: Artificial Sequence
110 <220> FEATURE:
111 <223> OTHER INFORMATION: single-stranded restriction fragment overhang
112 <400> SEQUENCE: 3
113      tgaga                                                         5
115 <210> SEQ ID NO: 4
116 <211> LENGTH: 33
117 <212> TYPE: DNA
118 <213> ORGANISM: Artificial Sequence
119 <220> FEATURE:
120 <223> OTHER INFORMATION: recognition site
121 <220> FEATURE:
122 <221> NAME/KEY: misc_feature
123 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 22, 23, 24,
124      25, 26, 27, 28, 29, 30, 31, 32, 33
125 <223> OTHER INFORMATION: n = A,T,C or G
126 <400> SEQUENCE: 4
W--> 127      nnnnnnnnnn acnnnnngtay cnnnnnnnnnn nnn                      33
129 <210> SEQ ID NO: 5
130 <211> LENGTH: 33
131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
133 <220> FEATURE:
134 <223> OTHER INFORMATION: recognition site
135 <220> FEATURE:
136 <221> NAME/KEY: misc_feature
137 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19, 18.
138      20, 21, 27, 28, 29, 30, 31, 32, 33
139 <223> OTHER INFORMATION: n = A,T,C or G
140 <400> SEQUENCE: 5
W--> 141      nnnnnnnnnn acnnnnngtay cnnnnnnnnnn nnn                      33
143 <210> SEQ ID NO: 6
144 <211> LENGTH: 10
145 <212> TYPE: DNA
146 <213> ORGANISM: Artificial Sequence
147 <220> FEATURE:

```

Handwritten notes:
 No "n"s at these locations (see below)
 "A" is at location 18.
 22 (see below)
 23, 24, 25, 26
 16, there is an "n"
 n's are at these locations

RAW SEQUENCE LISTING

DATE: 11/04/2003

PATENT APPLICATION: US/09/804,481A

TIME: 13:57:42

Input Set : N:\Crf4\11032003\I804481A.raw

Output Set: N:\CRF4\11042003\I804481A.raw

```

148 <223> OTHER INFORMATION: modification fragment
149 <400> SEQUENCE: 6
150     cacaacaca                                     10
152 <210> SEQ ID NO: 7
153 <211> LENGTH: 12
154 <212> TYPE: DNA
155 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: modification fragment
158 <400> SEQUENCE: 7
159     tccacaaaca ca                                12
161 <210> SEQ ID NO: 8
162 <211> LENGTH: 15
163 <212> TYPE: DNA
164 <213> ORGANISM: Artificial Sequence
165 <220> FEATURE:
166 <223> OTHER INFORMATION: modification fragment
167 <400> SEQUENCE: 8
168     tcgtccacaa acaca                             15
170 <210> SEQ ID NO: 9
171 <211> LENGTH: 12
172 <212> TYPE: DNA
173 <213> ORGANISM: Artificial Sequence
174 <220> FEATURE:
175 <223> OTHER INFORMATION: modification fragment
176 <400> SEQUENCE: 9
177     cacaacaca ac                                  12
179 <210> SEQ ID NO: 10
180 <211> LENGTH: 10
181 <212> TYPE: DNA
182 <213> ORGANISM: Artificial Sequence
183 <220> FEATURE:
184 <223> OTHER INFORMATION: modification fragment
185 <400> SEQUENCE: 10
186     cacaacacg                                     10
188 <210> SEQ ID NO: 11
189 <211> LENGTH: 59
190 <212> TYPE: DNA
191 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: vector construct
194 <400> SEQUENCE: 11
195     ggcccaagat ctcaagggcc cataacatgt gtaccatcga ttgcagggga gataccatg  59

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 11/04/2003
PATENT APPLICATION: US/09/804,481A TIME: 13:57:43

Input Set : N:\Crf4\11032003\I804481A.raw
Output Set: N:\CRF4\11042003\I804481A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 1,2,3,4,5,6,7,8,9,10,13,14,15,16,22,23,24,25,26,27,28,29,30

Seq#:4; N Pos. 31,32,33

Seq#:5; N Pos. 1,2,3,4,5,6,7,13,14,15,16,19,20,21,22,23,24,25,26,27,28,29

Seq#:5; N Pos. 30,31,32,33

VERIFICATION SUMMARY

DATE: 11/04/2003

PATENT APPLICATION: US/09/804,481A

TIME: 13:57:43

Input Set : N:\Crf4\11032003\I804481A.raw

Output Set: N:\CRF4\11042003\I804481A.raw

L:6 M:270 C: Current Application Number differs, Wrong Format
L:7 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:127 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:141 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0